The Study of the Effect of 5C Factors on the Credit Risk of Natural Customer of Refah e Kargaran Bank and Credit assessment

Farideh Mohammadi1*, Zadollah Fathi2
1Department of Business Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran; 2Assistant Professor, Faculty of Management, Central Tehran branch, Islamic Azad University, Tehran, Iran
*E-mail: farideh49_mohammadi@yahoo.com

Abstract
The subject of rating and recognition of banking customer risk is a profound subject that includes the most qualitative indicators to most quantitative cases. Many studies, by determining effective reasons and causes in this domain, have tried to tag well or badly on the customers and be sure of their choice before entering the customer to the bank. This study tries to recognize standards with priority by using three integrated approaches: qualitative based on Decision theory, quantitative based on statistics, and quantitative based on artificial intelligence. This study investigates some banking customers and adjusts its parameters and shows that this integrated approach, which in first step, studies the expected characteristics and recognizes them in a qualitative usage and then uses the artificial intelligence to identify behaviors, has a great value and can perform it with little error percentage.

Keywords: Customer Risk, Artificial Intelligence, Multi-Standard Decision, Integrated Approach

Introduction
Assessment decisions of credit risk are very important for financial institutions because of risks related to inappropriate credit decisions. Credit scoring attracts much attention because the credit industry benefits from money current improvement, credit collections, and reduce the possible risk. The goal of scoring models is the credit allocation to applicants in good group to Grant loans again or in bad group that means the possibility of not giving loan. Then, credit scoring issues work in more public domain than classification issues (Shavalpour & Ashari, 2013).

Creating an integrated and coordinated system can be the most important factor in risk management of bank to determine the credit risk and systematic probability of not amortization of original and profit of payment credits from applicants during request and this will be obtained just with having comprehensive information about credits customers (Espinoza and Prasad, 2011).

In banking industry, one of the most important issues that always credit policy makers must consider, is the credit risk management. The management and control of this risk, credit rating systems of customers is a deniable necessity. It is obvious that using such system helps customers to favor choosing and in addition to control of credit risk, improves the efficiency level of granting the banking facilities (Darabi & Molaie, 2011).

Previous studies often used statistic approaches such as Logit and Probit regression and Audit analysis approach to score and rate customers, but in recent years with developing the models based on artificial intelligence and innovative models, many studies have been done in using of these methods for credit scoring and rating models. In internal research, until now, the ants colony optimization method in rating and optimizing of Neural Networks in Multi-attribute decision making have not been used and this approach is completely new and innovative inside the country.
Also, the hybrid model of Fuzzy neural Network, ant colony, and TOPSIS are completely new and do not have any similarity in internal and external researches.

Then, these questions are raised that what are the financial and non-financial indicators about customers' validation? Can we use the hybrid model of Fuzzy neural Network and ant colony and TOPSIS method to credit rating of natural customers of banks?

According to these issues, the main purposes of this study are as follows:

- Ideal purpose: preparing model that determine the credit risk of natural customers of banks and estimate the probability of facilities default by them.
- General purpose: customers rating and determine the facilities ceiling and payable credit to them by bank.
- Practical purpose: preparing the practical model based on statistics of natural customers of branches of Refah bank of Tehran and Present rating method of bank.

According to what above said, the assumptions of this study are as follows:

- We can categorize the good customers by some indicators.
- Data analysis methods based on neural networks can help to repeat and understand this behavior. In addition, using the artificial intelligence algorithm such as ants colony in combining with Multi-attribute decision making can be effective about this operation and offer more suitable answers.

Methods of facilities maintenance in banking system of country

Banks claims including Rials claims such as granting facilities to legal and natural individuals, customers' debts in temporary debtor account, the customers debt about documentary credits and paid guarantees and paid term bills of exchange and advance payment about purchase of transactions assets, purchased assets about contracts, commodities of forward transactions, work in progress of reward, purchase debt and currency claims including facilities from currency reserves accounts and all the facilities which recipient committed to refund them. These claims are stored in term of approval letter of ministers in current claims category and non-current claims (past maturity claims, deferred claims, doubtful account, bad debts) in banking system.

The purpose of development of measurement system and management of credit risk

The purpose of this program is to plan a mechanism for measurement and management of bank credit risk. Based on this mechanism, the processes of bank about granting facilities are reviewed and if possible, alternative processes are suggested to improve and speed up the facilities payments for customers. By such mechanism, valid customers are distinguished from invalid customers and the bank resources flow toward applicants that have more competencies to receive loan. Adopt of Very conservative policies against credit decisions reduce the credit risk of bank (Fallah Shams, Mir Feiz, & Rashno, 2008).

The credit scoring standard of system 5C

In model 5C, we use the 5 main standards to credit evaluate customers: character, capacity, capital, condition, and collateral.

Introduction to Refah Kargaran Bank

Refah Kargaran bank was formed in August 1960. Its first branch was launched in Tehran and Esfahan in April 1961. For implementing note of article 39 of budget law in 1959 and article 38 of Social Insurance Agency, Refah bank was registered to investment and operation of Workers' insurance funds with purpose of help to provide welfare and facilities of working class in August, 18, 1960 and begun to operate in March, 26, 1961. Today, Refah Bank is known as a commercial
bank and has been active more than 50 years in this domain. It is owner of 1128 branches that its products are retail banking, Investment Banking, investment management, insurance management, and stoke broker and total assets are estimated approximately 121000 Billion Rials.

**Literature review**

Desai, Crook, and Overstreet (1996) performed credit scoring models with artificial neural networks on a data collection of 1962 customers that is taken from three charity institutions. Among said models, the efficiency of artificial neural network was better than others, especially about Forecast Bad Credit. Thomas (2000) investigated consumer credit risk models and pointed out its similarity with some company credit risk models.

Chang (2003) used three refuted credit applications in conditional reception group in transfer process and found that models which work on the credit data by CART and MARS are more successful than those use the separator Linear analysis, regression logical analysis, and artificial neural network.

**Internal studies**

Heidari, Zavarian, and Nourbakhsh (2010) studied the effect of Macro-economic shocks on the banks deferred claims in time period of 2000-2008. For this purpose, they used the ARDL (autoregressive distributed lag) model in first step.

Bahri (2009) used statistical models to evaluate bank deferred in function accounting area based on limitation theory. The results of research lead to identify Factors affecting deferred expenses and finally rank these factors in framework of function accounting of system and a related pattern was suggested. Hashemi (1998) studied the reasons and causes of deferred claims and past maturity of Maskan Bank in time period 1986-1997. According to this study, it was confirmed that variables of Market interest rate and also the difference between Loan interest rate plus Penalty rate and unorganized money Market interest rate, affect on the claims. All these research show the alignment of macro-economic variables with banks profits (or loss).

Keshavarzi, Mohtashami, and Salami (2011) identified Distinguishing factors between low-risk costumers and risky costumers by statistic and information related to 6000 granted facilities to legal costumers in 1991-2004 and with the help of some techniques such as Diagnostic Analysis technique. These results have strengthened this hypothesis that among factors related to credit characteristics, there are four factors including Reputation of company management, the number of bank accounts, reliable sponsor, the number of dud cheque, that have most important to distinguish between punctual and unreliable facilities Recipients. Among the variables related to facilities, most important distinguishing factors are: Long-term or medium-term loan, lead time, interest-free loan, leasing, the number of installments, and among the variables related to the nature of companies, most important distinguishing factors are: Activity in the field of rural cooperative, activities in some Certain provinces, and Regional drought.

**Research method**

Present study is a descriptive survey research. The statistical society is total customers of Refah Bank that have had financial and credit transactions up to now. Also, to survey, we use the some experts of bank and University. We study approximately 200 customers and build model based on 70% of samples. A number of samples (60 individuals) are used to test the reliability and performance of the model as witness data.

Since our research method is a hybrid method and each of parts needs different knowledge, then we will point out each part separately. In first phase, the method is based on data analysis and its axis is the polarity of customers data and their important indicators. The research tool is

Openly accessible at [http://www.european-science.com](http://www.european-science.com)
questionnaire. In this questionnaire, we try to know the experts and specialist opinion regarding standards and indicators to determine the final indicators. Experts are University professors, bank managers, and Qualified Auditors.

To implement this model, MATLAB software was used. To analyze data by regression, correlation, and Factorial analysis LISERL and SPSS software were used. To implement and analyze of TOPSIS, EXCEL software was used.

To study the variables and measure outcome data, the following Conceptual Model is presented:

![Conceptual Model of research](image)

**Figure 1: Conceptual Model of research**

It is obvious that we must distinguish between Natural and legal persons and in term of it, we assume and test different indicators that are results of 5C model to identify effective factors. Beside the data analysis, results of these tests will lead to obtain the customer's credit and rating.

**Data analysis**

As we see in age chart, the age distribution of society is normal and we can see the Employment situation. There is a population density between the ages of thirty. Also, in the age of the forties, the population is high. The following figure is drawn in the scale of 1 to 55. It means that, for example, in age of 17 (total record is 55000) the number of people is $55 \times 4$, namely 220 persons. For simplicity, the numbers are divided by 55 in figure 2.

![The distribution of age among the population](image)

**Figure 2: The distribution of age among the population**
As it is clear from the above figures, some age groups have the average equal to their age fraction of the minimum age to open an account. However, in some groups, there is a meaningful difference that it can be attributed to various reasons such as having account in other banks.

As it is clear from figure 4, six effective factors making non-credit worthy including delay in a month payment, delay in two month payment, delay in six month payment, discredit, delayed refund, and deferred claim increase with reduction of history of interaction with the bank. In simple terms, whatever these interactions are lower, its probability increases. The blue figure of discredit has the most slope.
As shown in these figures, we observe changes in short term service (green), facilities (red), and executive services (blue) with increasing the history of interactions. In short term services and bank facilities, the peak of receive service is 38 years and in executive services, the peak is 10-12 years. There is a big difference between these two services.

By using TOPSIS technique and Expert Choice software, the results of weight of each factor or sub-factors were specified as it can be observed in table 1.

### Table 1: Results of weight of each factor or sub-factors

<table>
<thead>
<tr>
<th>Weight</th>
<th>Factor</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03</td>
<td>Age of person/company</td>
<td>Characteristics</td>
</tr>
<tr>
<td>0.7</td>
<td>Number of referring to bank</td>
<td></td>
</tr>
<tr>
<td>0.48</td>
<td>Number of migration to other banks</td>
<td></td>
</tr>
<tr>
<td>0.39</td>
<td>Number of dependent family members/staff</td>
<td></td>
</tr>
<tr>
<td>0.18</td>
<td>Criminal record/legal rulings</td>
<td>Characteristics</td>
</tr>
<tr>
<td>0.11</td>
<td>Number of account in other banks</td>
<td></td>
</tr>
<tr>
<td>0.56</td>
<td>History of using bank services</td>
<td></td>
</tr>
<tr>
<td>0.97</td>
<td>History of using bank facilities</td>
<td></td>
</tr>
<tr>
<td>0.07</td>
<td>History of lack of timely payment to the bank</td>
<td></td>
</tr>
<tr>
<td>0.66</td>
<td>Amount of turnover in each account</td>
<td></td>
</tr>
<tr>
<td>0.36</td>
<td>Total amount of turnover</td>
<td></td>
</tr>
<tr>
<td>0.46</td>
<td>Kind of account</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>History of discredit in this bank</td>
<td></td>
</tr>
<tr>
<td>0.46</td>
<td>History of discredit in all banks</td>
<td></td>
</tr>
<tr>
<td>0.09</td>
<td>Amount of received facilities in different branches</td>
<td></td>
</tr>
<tr>
<td>0.51</td>
<td>Business nature</td>
<td></td>
</tr>
<tr>
<td>0.54</td>
<td>Kind of business (personal, legal)</td>
<td></td>
</tr>
<tr>
<td>0.67</td>
<td>Type of company stock (personal, public, private)</td>
<td>Condition</td>
</tr>
<tr>
<td>0.27</td>
<td>Kind of product or service</td>
<td></td>
</tr>
<tr>
<td>0.82</td>
<td>Level of being hi-tech in industry</td>
<td></td>
</tr>
<tr>
<td>0.29</td>
<td>Saturation level of mentioned industry</td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>Risk of project failure/business companies</td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>State support of this industry</td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>Trade restriction of this industry</td>
<td></td>
</tr>
<tr>
<td>0.72</td>
<td>Positive and suitable financial ratio of company</td>
<td></td>
</tr>
<tr>
<td>0.33</td>
<td>Credit provided by other banks in this industry</td>
<td></td>
</tr>
<tr>
<td>0.27</td>
<td>Existence of appropriate corporate governance</td>
<td></td>
</tr>
<tr>
<td>0.32</td>
<td>Last guarantees in received loan</td>
<td></td>
</tr>
<tr>
<td>0.58</td>
<td>Available guarantees for present loan</td>
<td></td>
</tr>
<tr>
<td>0.93</td>
<td>Banned guarantees</td>
<td></td>
</tr>
<tr>
<td>0.45</td>
<td>Expropriated guarantees</td>
<td></td>
</tr>
</tbody>
</table>

Based on obtained weights and with choosing several different combinations such as optional combination based on Pareto principle, the results were studied. 30% of data were used for system education, 40% for study and correction, and 30% for finding the credits. Figure 6 shows the error due to using various combinations.
To choose above selections, first, based on Pareto principle (20-80 principle), 20% of factors that could create 80% of weight, were extracted. Then, with random fraction of some factors of combinations 5, 6, 7 are obtained and by random adding of some factors to Pareto combination, the combinations 1, 2, and 3 were obtained. Also, figure 6 shows that reduction of some factors can increase errors. On the other hand, by increasing factors, error may reduce but may cause errors in future estimation.

Table 2 shows the validation process of nervous-fuzzy inference system. The system having best values mentioned in table is selected as an optimized system.

**Table 2: Validation method of adaptive nervous-fuzzy system**

<table>
<thead>
<tr>
<th>Optimized range</th>
<th>Value</th>
<th>Equivalence</th>
<th>Validation criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>The less the better</td>
<td>Zero to infinite</td>
<td>SSE</td>
<td>Sum of error squares</td>
</tr>
<tr>
<td>The less the better (less than 0.1 is excellent)</td>
<td>Zero to infinite</td>
<td>RMSE</td>
<td>Square root mean of error</td>
</tr>
<tr>
<td>Between 0.6-0.8 is suitable, 0.8-1 is excellent</td>
<td>Zero to one</td>
<td>R2</td>
<td>Determination coefficient</td>
</tr>
<tr>
<td>Between 0.6-0.8 is suitable, 0.8-1 is excellent</td>
<td>Zero to one</td>
<td>Adjusted-R2</td>
<td>Corrected determination coefficient</td>
</tr>
<tr>
<td>In comparing the two systems, the less the better</td>
<td>Zero to infinite</td>
<td>MSE</td>
<td>Average of error squares</td>
</tr>
<tr>
<td>In comparing the two systems, the less the better</td>
<td>Zero to infinite</td>
<td>MAE</td>
<td>Absolute mean error</td>
</tr>
</tbody>
</table>

MAE, MSE, and RMSE indicators were used to identify accuracy level of system and the closeness of estimates to the true values. About designed system of this study, whatever presented value for these indicators is less in a design and training method, system designed by this method is more efficient and have close estimate to the true values. Indicators of coefficient of determination and corrected coefficient of determination, also, are used for accuracy level of system output to estimate true values.

To determine three first validation criteria (RMSE, R2, Adjusted- R2), program code or curve fitting toolbox was used. In this Figure, the system output is compared with training data output and validation data output. Also, the software code was used to calculate two criteria MSE and MAE. Among presented indicators, the RMSE is the most important, calculated by software.
automatically and with training. Thus, this indicator is the most important to identify the efficient system. Besides RMSE indicator, the A-R2 indicator, also, is a suitable criterion to identify the power of system for estimating output data. Usually, both these indicators together, present system power with accurate and devoid of error estimate.

Figure 7: True values and values calculated by system for training data and validation data.

As seen in above Figure, algorithm (with the least deviation level from the original variables, which specify by green color) could create predicted values (violet color) as well.

Figure 8: Fitted regression graph on the training and validation data

As it is clear from figure 8, the regression line confirms that data placed on both sides of the regression line equally and observations are not much deviation of the line.

Goodness of fit:
SSE: 0.1318
R-square: 0.9758
Adjusted R-square: 0.9757
RMSE: 0.02139

Figure 9: Validation values presented by training data toolbox

Openly accessible at http://www.european-science.com
The above mentioned validation criteria are calculated and compared with training algorithm for each design method. In terms of design and education, the best system has following indicators:

- It has a Lower RMSE for training and validation data.
- The difference between RMSE values is minimized for training and validation data.
- It has the higher A-R2. This indicator covers values related to R2.

Finally, the adjusted values for system training are the same and if it needs value improvement for one of the methods, that value presented in front of the related figure by mentioned method. We can rearrange values to achieve the best outcome of system training.

The effect of using combined optimization of ants colony

In this study, each of ants follows an optimized rout so that with scattering different pheromones, the weight of each of possible is between different layers and nerves. With creating this layer that also implemented in MATLAB R2010a software, we can expect that necessary input is formed to product adoptive neural- fuzzy network arcs.

Although this method can have special position theoretically, but because of existence of multilayers and polarity of nerves, by using the 100 ants and 100 repetitions, the answer cannot be expected to be searched well and obtain answer better than adoptive neural- fuzzy network answer.

As it is observed in figure 10, first in ten consecutive tests, the operation of adoptive neural-fuzzy network was investigated without using the method of ants colony and then the operation was reviewed again with using this method. The results show that in 2 times of 10 times (20%), this method could obtain more suitable operation than run the main method alone.

![Figure 10: Comparison between using the adoptive neural- fuzzy network alone and in combination with ants-colony](http://www.european-science.com)

Discussion and Conclusion

In this study, the adoptive neural- fuzzy network technique was sued to predict the amount of facilities dishonor in banking industry. The preference of this study to other similar ones is the combination of three methods that although uses the different views such as multi- criteria decision, data analysis to extract knowledge, and Optimization algorithm, but can cover weak points with using the suitable approach of method benefits. The findings show that we can use this method to

Openly accessible at [http://www.european-science.com](http://www.european-science.com)
offer appropriate estimations for future and evaluate present situation. The result of these outputs is that by using a large number of data in this system, its operation has been done by following approach:

First, the nervous network was designed, then about 30% of data were used as test data and 60% of them were used as the test and the system was rearranged again and the system behavior was evaluated again. When the operation of network reached to 95%, the system was prepared. Finally, with help of 10% of rest data, the behavior of system was evaluated. In the end, the number 89% showed this fact that 89% of times this system can offer true estimate about future cost and income.

Then, we can describe this system as a suitable tool to support the decision at approximately 90% confidence level that can be effective in determination of facilities discredit in Iran banking system by predicting the cost and income of each person. In other words, we can know this tool effective not only for account owners in different banks, but also for all the new applicants that do not have any account records in banks.

Therefore, regarding the first research question, we can conclude that the most important variables can be categorized in three main forms: personal, financial, and business. Based on final weights of arcs of neural network, we can test most important indicators.

To answer the second question, we claim that we can ensure the effectiveness of this method with 90% confidence level. The mean reason is the data-driven method that, unlike other statistical and mathematical methods, removes assumptions that cause simplification and relies on facts.

The effectiveness of method can be converted to nervous network by passing steps and considering factors such as use much data, suitable form of neural network, and use of suitable training methods and can be considered as a supportive tool of appropriate decision. This method can be distinguished from the present methods that have been limited to data without disturbance and despite of turbulent or lost data, we can show a suitable operation.

Finally, the most identified factors considered in this study include the previous activities, records cooperation, records of deferred loan, amount of loan, the average of account, assets and account turnover, current ratio, cash ratio, the ratio of total debt to assets, turnover ratio of funds, ratio of debt collection, and ratio of current debt to equity. In this study, because of the especial economic situations, we cannot consider previous records as very important factors.

References

Openly accessible at http://www.european-science.com


